

学校编码: 10384

分类号\_\_\_\_\_密级\_\_\_\_\_

学号: X2008221020

UDC\_\_\_\_\_

厦 门 大 学

工 程 硕 士 学 位 论 文

# 基于屏幕流特点的高效视频压缩算法研究

Efficient flow characteristics of screen-based video  
compression algorithm

贾 燕

校内指导教师: 郭 峰 讲师

校外指导教师: 谢 孟 荣 教授

专 业 名 称: 计 算 机 技 术

论文提交日期: 2012 年 月

论文答辩时间: 2012 年 月

学位授予日期: 2012 年 月

答辩委员会主席: \_\_\_\_\_

评 阅 人: \_\_\_\_\_

2012 年 月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（        ） 1. 经厦门大学保密委员会审查核定的保密学位论文，  
于        年        月        日解密，解密后适用上述授权。

（        ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年        月        日

## 摘要

随着计算机和网络通信技术的迅速发展, 计算机多媒体和网络传输技术在教育领域中显示出了广阔的应用前景, 多媒体技术与网络技术的结合正在改变传统的教学方式。网络教学软件往往需要进行屏幕广播及屏幕监控, 屏幕共享技术越来越得到人们的重视。屏幕数据作为一个必不可少的“源”, 其重要性已经超过了视频数据。屏幕共享技术中最核心的问题就是对屏幕数据的编码压缩, 当屏幕分辨率为  $800 \times 600$ , 颜色为真彩色(24 位)时, 屏幕图像的数据量为 1.44M 字节, 1 秒钟传播的数据量就更大, 因此, 在当前网络条件下实时屏幕共享难以实现。为在局域网中满足教学软件的实时性要求, 必须有压缩率高、压缩速度快的压缩算法。

本文针对现有技术的不足和缺陷, 并针对屏幕流的特点在深入研究和在本文探讨视频压缩技术以及屏幕流特点的基础上, 提出了一种有效提高实时屏幕流压缩处理效率的视频压缩算法。依据对于不同的数据采用不同类型的编码方式, 压缩的效果会更好这一思路, 也借鉴了 MPEG-4 的一些关键技术, 比如视频对象提取技术和视频编码可分级技术等, 来完成基于视频流特点的压缩算法的研究, 研究这个算法有以下的基本步骤:

- 1、利用 HOOK 技术进行屏幕图像的采样与获取。
- 2、基于运动补偿技术的帧间压缩。
- 3、基于内容的帧内无损压缩与有损压缩的结合。
- 4、色深和分辨率两方面进行压缩。
- 5、保存文件。

在本算法研究中我们对帧内采用有损编码与无损编码相结合, 对图像采用有损编码, 而对窗口, 文字等简单显示, 采用无损编码, 使帧内数据量更小(空间冗余), 适合屏幕上的各种图像。帧间应用简单的运动补偿技术, 压缩时间上的冗余, 使帧间数据更小。一般的编码方法只考虑前后两帧的时间冗余。本算法研究通过建立运动补偿库的方法可以更有效地压缩时间上的冗余。鼠标与屏幕分别编码, 使得鼠标运动时只记下鼠标位置, 屏幕数据并不变化。

关键词：屏幕流；视频压缩；数据冗余；帧间预测

厦门大学博士论文摘要库

## Abstract

With the computer and network communication technologies, the rapid development of computer multimedia and network transmission technology in the field of education has demonstrated a broad application prospects, multimedia technology and networking technology is changing the traditional teaching methods. Online teaching software will often need to screen, radio and screen monitor, which needs to spread a lot of screen image. Screen sharing technology has increasingly gained attention. Screen data as an indispensable "source", and its importance over the video data. Screen sharing technology, the core of the problem is that the encoding of data compression on the screen when the screen resolution of  $800 \times 600$ , color is true color (24 bit), the screen image data of 1.44M bytes, 1 second the greater the amount of data transmitted, so in the current network conditions, real-time screen sharing difficult to achieve. In the local area network to meet the educational software for real-time requirements, there must be a high compression ratio, compression speed and compression algorithm.

In this paper, the shortcomings and deficiencies of existing technologies, and flow characteristics for the screen. In-depth study and explore in this article and video compression technology based on the characteristics of the screen flow, we proposed an efficient flow of compression to improve the efficiency of real-time screen video compression algorithm. Based on different data for different types of encoding, compression would be better on this idea, MPEG-4 also draws some of the key technologies such as video object extraction technology and scalable video coding technology, to complete based on video stream compression algorithm characteristics, study this algorithm has the following basic steps:

- 1, the use of technology for the screen image HOOK sampling and access to
- 2, based on motion compensation inter-frame compression technology
- 3, content-based lossless and lossy compression frame combination
- 4, color depth and resolution of both compression
- 5, save the file

In this algorithm we used for Intra lossy coding and lossless coding by combining the use of lossy image coding, while the windows, and other simple text display, using lossless coding, so that a smaller amount of data frame (spatial redundancy), for a variety of images on the screen. Application of a simple inter-frame motion compensation technology, compressed time redundancy, so that smaller data frames. General encoding method only considers the time before and after the two redundant. The motion compensation algorithm library through the establishment of more efficient compression method can be time redundancy. Mouse and screen were coded, making the mouse movement only when the mouse down position, the screen does not change the data.

Key Words: Screen Streaming; Video Compression; Data Redundancy; Inter-frame prediction.

## 目录

摘要.....	I
Abstract.....	III
<b>第一章 绪论</b> .....	<b>1</b>
1.1 引言.....	1
1.2 数据压缩的发展.....	2
1.2.1 数据压缩的起源.....	2
1.2.2 真正的编码.....	3
1.2.3 LZ 系列算法.....	4
1.3 现代压缩技术.....	4
1.3.1 图像领域的压缩技术.....	5
1.3.2 视频领域的压缩技术.....	5
1.4 视频压缩技术的国内国外研究现状.....	6
1.5 目前的屏幕共享技术.....	7
1.6 本文的工作.....	8
<b>第二章 视频压缩技术研究</b> .....	<b>10</b>
2.1 视频压缩原理.....	10
2.2 主要的视频压缩技术.....	12
2.2.1 预测编码.....	12
2.2.2 变换编码.....	14
2.2.3 统计编码(熵编码).....	15
2.3 基于混合编码框架的视频编码器.....	16
2.4 视频图像压缩的标准化.....	17
2.4.1 H. 261 标准.....	17
2.4.2 MPEG-1 标准.....	17
2.4.3 H. 263 标准.....	18
2.4.4 MPEG-2 标准.....	18
2.4.5 MPEG-4 标准.....	19
2.4.6 H. 264 标准.....	19
2.5 视频图像压缩中关键技术-运动估计.....	20



2.6 视频质量评估 .....	20
2.6.1 主观评价 .....	21
2.6.2 客观测量 .....	21
<b>第三章 屏幕流特点 .....</b>	<b>23</b>
3.1 屏幕图像的序列性 .....	23
3.2 屏幕图像的空间相关性 .....	23
3.3 屏幕图像的时间相关性 .....	24
3.4 屏幕图像的色彩分布 .....	24
3.5 屏幕图像清晰度的要求 .....	24
3.6 拷屏图像的属性与显示器的设置有关 .....	25
<b>第四章 基于屏幕流特点的压缩算法 .....</b>	<b>27</b>
4.1 借鉴 MPEG-4 .....	27
4.2 设计思想 .....	28
4.3 算法具体步骤 .....	28
4.3.1 利用 HOOK 技术进行屏幕图像的采样与获取 .....	28
4.3.2 基于运动补偿技术的帧间压缩 .....	36
4.3.3 基于内容的帧内无损压缩与有损压缩的结合 .....	42
4.3.4 色深和分辨率两方面进行压缩 .....	45
4.3.5 保存文件 .....	46
4.4 试验数据 .....	47
<b>第五章 总结与展望 .....</b>	<b>49</b>
5.1 总结 .....	49
5.2 展望 .....	50
<b>参考文献 .....</b>	<b>52</b>
<b>致 谢 .....</b>	<b>54</b>

## Table of Contents

<b>Abstract in Chinese</b> .....	<b>I</b>
<b>Abstract in English</b> .....	<b>III</b>
<b>Chapter 1 Introduction</b> .....	<b>1</b>
<b>1.1 Introduction</b> .....	<b>1</b>
<b>1.2 Development of data compression</b> .....	<b>2</b>
1.2.1 The origin of the data compression.....	2
1.2.2 The real coding.....	3
1.2.3 LZ series algorithm.....	4
<b>1.3 Modern compression technology</b> .....	<b>4</b>
1.3.1 The field of image compression technology .....	5
1.3.2 The field of video compression technology .....	5
<b>1.4 Video compression technology,domestic and foreign Research</b> .....	<b>6</b>
<b>1.5 The current screen-sharing technology</b> .....	<b>7</b>
<b>1.6 The work of this paper</b> .....	<b>8</b>
<b>Chapter 2 Video Compression Technology</b> .....	<b>10</b>
<b>2.1 Video compression principle</b> .....	<b>10</b>
<b>2.2 Key video compression technology</b> .....	<b>12</b>
2.2.1 Predictive coding .....	12
2.2.2 Transform coding .....	14
2.2.3 Statistical coding(entropy coding).....	15
<b>2.3 Hybrid coding framework based on the video encoder</b> .....	<b>16</b>
<b>2.4 Video compression standard</b> .....	<b>17</b>
2.4.1 H.261 standard.....	17
2.4.2 MPEG-1 standard .....	17
2.4.3 H.263 standard .....	18
2.4.4 MPEG-2 standard .....	18
2.4.5 MPEG-4 standard .....	19
2.4.6 H.264standard.....	19
<b>2.5 Key video compression technology-Motion Estimation</b> .....	<b>20</b>
<b>2.6 Video quality assessment</b> .....	<b>20</b>
2.6.1 Subjective evaluation .....	21
2.6.2 Objective measurement .....	21
<b>Chapter 3 Screen flow characteristics</b> .....	<b>23</b>
<b>3.1 The sequence of screen images</b> .....	<b>23</b>

3.2 Screen images of the spatial correlation.....	23
3.3 Screen images of time-related.....	24
3.4 The color distribution of the screen image.....	24
3.5 Screen image resolution requirements.....	24
3.6 Copy screen image attributes and set the display .....	25
<b>Chapter 4 Flow characteristics compression algorithm .....</b>	<b>27</b>
4.1 Learn MPEG-4 .....	27
4.2 Design idea .....	28
4.3 Algorithm specific steps.....	28
4.3.1 HOOK technology for the use of sampling and to obtain the screen image.....	28
4.3.2 Based on motion compensation inter-frame compression technology ..	36
4.3.3 Content-based lossless and lossy compression frame combination .....	42
4.3.4 Color depth and resolution of both compression.....	45
4.3.5 Save the file.....	46
4.4 Test Data .....	47
<b>Chapter 5 Summary and Outlook .....</b>	<b>49</b>
5.1 Summary.....	49
5.2 Outlook.....	50
<b>References.....</b>	<b>52</b>
<b>Acknowledgements.....</b>	<b>54</b>

1.1 引言.....	1
1.2 数据压缩的发展.....	2
1.2.1 数据压缩的起源.....	2
1.2.2 真正的编码.....	3
1.2.3 LZ 系列算法.....	4
1.3 现代压缩技术.....	4
1.3.1 图像领域的压缩技术 .....	5
1.3.2 视频领域的压缩技术 .....	5
1.4 视频压缩技术的国内国外研究现状.....	6
1.5 目前的屏幕共享技术.....	7
1.6 本文的工作.....	8
2.1 视频压缩原理.....	10
2.2 主要的视频压缩技术.....	12

2.2.1 预测编码.....	12
2.2.2 变换编码.....	14
2.2.3 统计编码(熵编码).....	15
<b>2.3 基于混合编码框架的视频编码器.....</b>	<b>16</b>
<b>2.4 视频图像压缩的标准化.....</b>	<b>17</b>
2.4.1 H.261 标准.....	17
2.4.2 MPEG-1 标准.....	17
2.4.3 H.263 标准.....	18
2.4.4 MPEG-2 标准.....	18
2.4.5 MPEG-4 标准.....	19
2.4.6 H.264 标准.....	19
<b>2.5 视频图像压缩中关键技术-运动估计.....</b>	<b>20</b>
<b>2.6 视频质量评估.....</b>	<b>20</b>
2.6.1 主观评价.....	21
2.6.2 客观测量.....	21
<b>第三章 屏幕流特点.....</b>	<b>23</b>
3.1 屏幕图像的序列性.....	23
3.2 屏幕图像的空间相关性.....	23
3.3 屏幕图像的时间相关性.....	24
3.4 屏幕图像的色彩分布.....	24
3.5 屏幕图像清晰度的要求.....	24
3.6 拷屏图像的属性与显示器的设置有关.....	25
<b>第四章 基于屏幕流特点压缩算法.....</b>	<b>27</b>
4.1 借鉴 MPEG-4.....	27
4.2 设计思想.....	28
4.3 算法具体步骤.....	28
4.3.1 利用 HOOK 技术进行屏幕图像的采样与获取.....	28
4.3.2 基于运动补偿技术的帧间压缩.....	36
4.3.3 基于内容的帧内无损压缩与有损压缩的结合.....	42
4.3.4 色深和分辨率两方面进行压缩.....	45
4.4 试验数据.....	47
<b>第五章 总结.....</b>	<b>49</b>

## 第一章 绪论

### 1.1 引言

1895 年,随着电影的诞生,视频信号第一次走进人类社会,随着电视机的发明和逐渐的普及,视频信号逐渐出现在家家户户。据统计,人类每天获取的信息中,视觉信息约占信息量的 70%左右,换句话说,视频信息在多媒体中占有重要地位。科学技术的迅猛发展,使数字技术得到更广泛的应用,这也使视频信号的存储、传输发生革命性的转变,但是如果想把信号从模拟转换到数字,所需要的原始视频信息的数据量是巨大的,比如说未经过压缩的电视视频信号一秒钟内的数据量就达到 216 兆比特,而高清晰度电视 HDTV 则在 1.2Gbit/s 以上,这是对存储空间的一个很大挑战,因此数字视频技术就算是多媒体应用的核心技术,也就是说假若可以解决视频数字化后数据量与数据存储媒体和通信容量小的矛盾,多媒体将得到更广泛的应用,而压缩就是解决这个问题的基本途径。因此,研究和开发新型而且更加有效的压缩编码方法,存储和传输压缩过的形式,则是最佳的解决途径<sup>[3,4,5]</sup>。

压缩数据量是视频编码的首要目标。压缩就是将原始的数据用更少比特数表达出来的过程。而将数字的视频序列精简到比单纯压缩还要少的比特数来表达,就是我们所说的视频压缩(视频编码)<sup>[1]</sup>。因为多媒体声音、文字、静止图像和动态视频图像等信息源有很强的相关性,换一种说法就是存在大量的冗余信息,而数据压缩就是通过除去冗余完成的,冗余可以看作是在恢复完整的数据时那些多余的信息。含有统计冗余的数据有很多种,我们可以使用无损压缩的方法去除,同时保证解码端恢复的数据是编码端数据的完整复制。然而,无损压缩这种方法的压缩效率并不高,只是在除去视频或者图像的冗余时能起到一定的作用。目前无损压缩堪称效率最高的图像压缩标准,但是这个压缩效率仍然达不到人们的要求。所以,要想得到比无损压缩更高的压缩效率只好采用有损压缩的方法。所谓有损压缩系统,是通过除去图像和视频中的主观冗余来达到进一步的数据压缩<sup>[2,4,5]</sup>。

主观冗余，是指虽然去掉该信息但不会对观看者感受到的视觉质量造成太大的影响。

大多数视频编码方法是指在减小空间时同时减少时间上的冗余来达到压缩的效果。视频图像在时域中，是沿时间轴方向的一个帧序列，其帧间图像有很强的相关性，主要表现在两帧间很多静止图像的数据是相同的。而在空间域上，相邻的像素点之间也有很强的关联性。运动补偿和运动估计是实现帧间编码的主要方法，利用帧间时间的相关性来缩小时间的冗余程度<sup>[2,3]</sup>。

## 1.2 数据压缩的发展

从比较严格的意义上来说，数据压缩最早来源于人们对概率的认识。如果在对文字信息进行编码的时候，将出现概率较高的字母赋予将较短的编码，将出现概率较低的字母赋予较长的编码，这样总的编码长度就能大大缩短。其实早在计算机出现之前，著名的莫尔斯电报码就已经验证过这一准则。在莫尔斯码表中，字母点划组合形成一一对应关系<sup>[6]</sup>，出现概率最高的字母 e 被编码为一个点“.”，而出现概率较低的字母 z 则被编码为“-...”。这样电码长度就被有效缩短了。

### 1.2.1 数据压缩的起源

被称为信息论之父的香农首次用数学语言阐述出概率与信息冗余度之间是有一定联系的，这也成为数据压缩的开始。在香农 1948 年发表的论文“通信的数学理论”中，他指出，任何信息都存在冗余，而且冗余大小与信息中每一个符号，无论是字母、数字或者是单词，他们的出现概率或者说不确定性有关。香农从热力学的概念中得到启发，把排除了冗余后的平均信息量称为“信息熵”，并给出了用于计算信息熵的数学表达式。信息熵为后来的数据压缩算法提供了理论上的基础，这篇论文后来也被誉为信息论的源头，香农成为信息论的鼻祖。从根本上来讲，消除信息中存在的冗余是数据压缩的唯一目的，而信息熵和相关定理正好可以利用数学方法来精确地描述出信息冗余程度<sup>[7,8]</sup>。我们能用信息熵公式计算信息编码的极限，

即在一定的概率模型下，无损压缩的编码长度不可能小于信息熵公式给出的结果。

香农在提出信息熵理论时，也提出了香农编码。这些早期的编码方法揭示了有关变长编码的基本规律，虽然能够看到一些压缩效果，但与真正实用的压缩算法还有很大的差距<sup>[8]</sup>。

### 1.2.2 真正的编码

1952 年，哈夫曼在他的论文“最小冗余度代码的构造方法”中提出的编码方法是第一个可实用编码。这种被称为哈夫曼编码的方法一直为后人称颂，以至现在许多《数据结构》教材在讨论二叉树时仍要提及这种方法。

哈夫曼编码具有效率高，运算速度快，实现方式灵活的特点，从 20 世纪 60 年代至今，在数据压缩领域得到了广泛的应用。例如，在早期 UNIX 系统上，有一个不太为现代人熟知的压缩程序 COMPACT，实际上使用的是哈夫曼 0 阶自适应编码。20 世纪 80 年代初，在 CP/M 和 DOS 系统上也出现了哈夫曼编码，具有代表性的程序叫 SQ。今天，在很多常用的压缩工具和算法中，我们仍然能看到哈夫曼编码的影子。不过哈夫曼编码所得的编码长度与信息熵计算结果是相近的，还不是真正的信息熵极限。所以在现代压缩技术中，哈夫曼只是被视作最终的编码手段，而非数据压缩算法的全部<sup>[6,7,8]</sup>。

1968 年左右，P. Elias 在香农与费诺的编码方法的基础上，构造出香农费诺-Elias 编码。J. Rissanen 顺着这个编码方法的设想，在 1976 年提出了算术编码——一种更逼近信息熵的极限的编码方法。1982 年 Rissanen 和 G. G. Langdon 一同对算术编码进行了改善。之后，人们在将算术编码和 J. G. Cleary 与 I. H. Witten 在 1984 年提出的部分匹配预测模型进一步整合之后，开发出了压缩效果近乎完美的算法<sup>[9]</sup>。

对于无损压缩而言，匹配预测模型与算术编码的整合，已经可以最大程度地逼近信息熵的极限。算术编码虽然可以获得最短的编码长度，但因为其本身的复杂性也导致了算术编码的任何具体实现在运行时速度相当的慢。即使在当今社会，摩尔定律广泛流传，CPU 以飞快的速度更新，算术编码程序的运行速度仍然难以满足日常应用的需求<sup>[9]</sup>。

### 1.2.3 LZ 系列算法

1977 年, Lempel 和 Ziv 发表了论文“顺序数据压缩的一个通用算法”, 论文中所讲述的算法就是后来被人们称为的 LZ77 算法。两人又于 1978 年发表该论文的后继之作“通过可变比率编码的独立序列的压缩”, 论文中所讲述的压缩算法就是后来的 LZ78 算法。1984 年 T.A. Welch 发表论文“高性能数据压缩技术”, 他在论文中描述了在 Sperry 研究中心的研究成果——一个 LZ78 法的变种, 即后来闻名一时的 LZW 算法。1990 年后, T. C. Bell 等人又陆陆续续提出了很多对 LZ 系列算法的改善<sup>[12]</sup>。

事实上, LZ 系列算法只是用巧妙的方法将字典技术应用于通用数据压缩领域。简而言之, 当你使用字典中的页码与行号找到文章中的单词时, 实际上, 你已经把握了 LZ 系列算法的真谛。虽然, 在表面上, 这种以字典模型为基础的思路和香农、哈夫曼这些人开创的统计学的方法有天壤之别, 但是, 它的效果同样可以很逼近信息熵极限。并且, 从理论上, 我们可以证明, LZ 系列算法的本质依然符合信息熵基本的规律。

在数据的压缩领域中, LZ 系列算法的凸显的优势很快就得到了体现, 运用 LZ 系列算法的工具软件的数量与日俱增。使用 LZW 算法的 compress 程序最早出现在 UNIX 系统中, 该程序也很快成为了 UNIX 世界的压缩标准。紧接着, MS-DOS 环境下的 ARC 程序, 以及 PKWare、PKARC 等仿制品也逐渐产生。在 20 世纪 80 年代, 最著名的压缩工具当属 LHarc 和 ARJ 这两种 LZ77 算法的杰出代表<sup>[10,11,12]</sup>。

## 1.3 现代压缩技术

在当代, 整个通用数据压缩领域几乎被 LZ77、LZ78 和 LZW 算法以及它们的各种变体垄断, 众所周知的 WINZIP、WINRAR、PKZIP、GZIP 等压缩工具以及 PNG、GIF 和 ZIP 等文件格式都得益于 LZ 系列算法, 甚至连某些加密文件格式, 如 PGP 等, 也把 LZ 系列算法作为数据压缩的标准。通用数据压缩中兼顾速度与压缩效果的难题基本通过 LZ 系列算法得到解决。我们可以从香农的信息中看到: 我们拥有越多的对信息的先验知识, 就可以把信息压缩得越小。也就是说, 假设压缩算法的设计目标不是任意



Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库